

AperTO - Archivio Istituzionale Open Access dell'Università di Torino

Gender at Work: Incentives and Self-Sorting

This is the author's manuscript

Original Citation:

Availability:

This version is available <http://hdl.handle.net/2318/1527109> since 2015-10-26T13:11:01Z

Published version:

DOI:10.1016/j.socec.2014.12.005

Terms of use:

Open Access

Anyone can freely access the full text of works made available as "Open Access". Works made available under a Creative Commons license can be used according to the terms and conditions of said license. Use of all other works requires consent of the right holder (author or publisher) if not exempted from copyright protection by the applicable law.

(Article begins on next page)



UNIVERSITÀ DEGLI STUDI DI TORINO

This is an author version of the contribution published on:

Journal of Behavioral and Experimental Economics (formerly *Journal of Socio-Economics*), 55(1):
10 – 18

The definitive version is available at:

<http://www.sciencedirect.com/science/article/pii/S2214804314001268>

GENDER AT WORK: INCENTIVES AND SELF-SORTING

Abstract

This paper analyses the relationship between workers' gender and monetary incentives in an experimental setting based on a double-tournament scheme. The participants must choose between a piece-rate payment or a performance prize. The results show that women fail to reveal their type, and are less sensitive than men to the monetary incentives of the tournament. In addition, the tournament scheme induces males, but not females, to signal their ability and to select the contract which is more profitable for them.

Keywords: gender; discrimination; incentives; work; experiment.

JEL Classification codes: C91, J16, J41

1. Introduction

The gender pay gap is a widespread and well known phenomenon (Castagnetti and Rosti, 2009). Generally people tend to explain it as a matter of discrimination *tout court*: since it is well-known that the most of societies are chauvinist, then the women's treatment is worse than men's *ceteris paribus*. Of course this can be (and in fact is) an explanation of the phenomenon; however there can be other reasons why it exists and persists. In this article I would like to present a different (although partial) explanation: I argue that the gender pay gap may originate from gender-specific preferences. Indeed, some indications supporting this claim may be found in some part of the extant experimental literature, which shows the existence of some behavioural and attitudinal differences between men and women with respect to competition. Since wages depend on individual performances in competitive environments, these differences may help explaining the observed gender gap. However, while on the one hand some studies find that women shy away from competition (Niederle and Vesterlund, 2007); on the other hand, other scholars do not find such an evidence (Ivanova-Stenzel and Kübler, 2011). However, the way people (workers in particular) react to competition and their attitudes towards it are likely (at least partially) responsible for wage differentials. If in competitive environments wages include a prize for good performances, then workers who engage more in competitive environments may earn more than workers who prefer less competitive contracts. Therefore, should the women shy away from competition, this phenomenon would (partially) explain the wage gender gap, also in absence of gender discrimination. Women would just forgo higher salaries in order to obtain the preferred "contract".

This paper employs a double tournament setting to study 1) whether men and women differ in their preferences for competition, 2) whether people who reveal a preference for competing in a tournament actually perform better than those who prefer a non-competitive framework, and 3) whether people who choose to play a tournament but end in a non-competitive setting perform better than those who reveal a distaste for competition. In order to investigate these three points, I

run an experiment in which the subjects must perform a boring task; the remuneration for the task is either piece-rate or based on the ranking in a tournament (as in Niederle and Vesterlund, 2007). People bid on which type of “contract” they desire to work under, by stating their preference in a sealed-envelope auction, and then they actually start to work (see Section 3 for further details).

The choice between two contracts, one of which prizes productivity more than the other, depends also on the ability of the individual. The more one feels to be skilled, the more he should prefer the prizing scheme. I assume that ability is private information of the workers, and that a potential employer cannot observe ability of applicants. A possible way to screen them and to disentangle the more and the less productive could be offering them the choice between two contracts. The first pays a piece-rate wage, while the second prizes the performance, making the workers play a sort of tournament: at the end, the best contestants will get a salary, which is higher than that they would have obtained under the piece-rate scheme. Assuming risk neutrality, to maximise the final wage, high-ability workers should therefore choose the second contract, while the others should choose the first. The individual choice may therefore be assumed as a sort of signalling of the “quality” of the worker: the employer elicits the applicants to reveal their ability. Of course, productivity prizes are also likely to foster the workers’ effort.

The results of the experiment reveal that women 1) do not perform significantly better in a competitive environment (whereas men do)¹, 2) are much less sensitive than men to the incentives of competition. Moreover, 3) the participants’ preferences for a given payment scheme are a signal (to a potential employer) of their job performance (although it is not possible to assess if this is due to ability, to effort or to both). These results answer also the questions raised before in this section. In particular, we can observe that women tend to prefer non-competitive to competitive work environments, while the opposite holds for men. Similarly, only the men, who declared to prefer competitive payment schemes perform better than the men who did not, while the women liking competition and those disliking it show the same performance. Finally, the women who chose the tournament and ended in the non-competitive environment did not perform

differently from the women who, disliking competition, obtained to work in non-competitive environments. For men, the opposite results holds: the men who like competition, but were assigned the non-competitive scheme, perform anyway better than the men who chose and obtained the non-competitive contract.

2. Related literature

Croson and Gneezy (2009) survey several empirical and experimental works to conclude that men and women have different preferences in several domains, one of which is competition. According to some scholars, women would prefer less competition than men do. Niederle and Vesterlund (2007) find two factors explaining why men tend to enter tournaments more often than women: firstly, men are more overconfident than women (see also Bengtsson et al., 2005) and, secondly, men are more likely to prefer a competitive work environment than women². In line with these results, also Kleijnans (2009) and Fletschener et al. (2010) find that women tend to “shy away” from competition. In particular, Fletschener et al. (2010) observe that women in Central Vietnam self-select in economic activities characterised by low returns to avoid competitive markets. In other words, this shows that women are willing to forgo higher wages to work under the preferred conditions. The experimental setting of Niederle and Vesterlund (2007) offers two payment schemes to the participants: these have to perform given (mathematical) tasks under either a non-competitive or a competitive (so-called “tournament”) rule. In the former case, they receive a piece-rate payment for each task solved; in the latter, only the best performer of each group gets paid a given sum for each correct computation. The unit payment under the tournament rule is thus much higher than the unit payment under the piece-rate scheme; as a consequence, high-ability players have an incentive to choose the tournament. In a different experiment by Schwierien and Weichselbaumer (2010), women perform significantly worse than men in a competitive environment.

Nevertheless, other studies present different results. Gneezy et al. (2003), Gneezy and Rustichini (2004) and Price (2008) observe that, when people are operating in mixed-gender groups, competition increases the performance of male subjects³, while that of the females stays the same; on the other hand, women's performance does indeed improve when the competitors are all female. These findings do not appear to hold when the competition is between teams rather than individuals. Ivanova-Stenzel and Kübler (2011) find that, when the competition is between same-gender groups, men perform significantly better than females⁴, but, again, when mixed-gender teams compete against each other no gender effect is detectable.⁵ Furthermore the authors observe that "the composition of the team has no significant effect on the performance of each gender for a given incentive scheme"⁶. In matrilineal societies women do not shy away from competition and show behaviours in line with males' in patriarchal societies (Gneezy et al., 2009 and Gong and Yang, 2012) Moreover women tend to be loss-averse (Brooks and Zank, 2005). Vandergift and Yavas (2009)⁷ show that while women initially perform significantly worse than men, later there is little gender-related difference in performance under certain conditions and when the competition involves the repetition of a task (game).

In the experimental setting presented in this paper, where the applicants to a job express their preferences over two different contracts, self-confidence (i.e. the self-valuation of own abilities) plays a crucial role. In particular, Santos-Pinto (2012) proposes a theoretical model, whose conclusion is that women will earn less than men if the former are less self-confident than the latter (i.e. women are underconfident, whereas men are overconfident). Empirical evidence (Bengtsson et al., 2005 and Niederle and Vesterlund, 2007) shows that this may occur. In such a context, men (women) may over-prefer the (non-)competitive setting simply as a result of their over(under)confidence. In other words, these differences in self-confidence may bias the signal conveyed by the applicant to the potential employer, resulting in a misallocation between the two contracts. This would cause gender wage differences and an aggregated loss of efficiency in the labour market, because of the bias in the signal.

The works summarised in the previous part of this section show that there is no conclusive evidence whether women prefer competitive workplaces more or less than men do. However, so far the experimental studies have observed how the genders react to competition, during and after the performance of the experimental tasks. I propose an experiment, where the attitude towards competition can be assessed both before and after the experimental work has started. Indeed the attitude towards competition may depend also on the perception the individual have about their relative performance during the experiment. In other words, subjects who think that they are performing well may have stronger preferences for competition than subjects whose performance is (perceived as) poor. For this reason, preferences expressed before starting the experimental task are not influenced by how the subjects perceive their relative performance during (or after) the completion of a task. This is because the preferences are observed before any feedback on performance is available.

3. Experimental design and procedure

Experiments are useful to isolate particular variables, for which a clean effect would be too difficult to estimate in a really noising setting (Levitt and List, 2007) or when a field experiment is not feasible (Levitt and List, 2009). Moreover: “While laboratory processes are simple in comparison to naturally occurring processes, they are real processes in the sense that real people participate for real and substantial profits and follow real rules in doing so. It is precisely because they are real that they are interesting.”⁸

The experiment involved a total of one hundred forty-six undergraduate students (sixty-nine males and seventy-seven females) who played a two-stage game. First they were explained the task: they would be asked to enter a list of fictitious names, identification numbers and exam results line-by-line into a computer database. The list to be copied was the same for all the participants. Payment would be made for each line (name, id number and mark) correctly copied into the pre-formatted table; mistakes would be signalled by the PC and would have to be

corrected before it would be possible to proceed. The participants were told that the task would last 45 minutes, after which the programme would automatically interrupt their work. Participants were given five minutes to practice before continuing with the experiment.

After the practice session, participants were presented with two possible remuneration schemes: tournament or piece-rate. Under the tournament scheme, payment would depend on performance ranking, according to the following guidelines: the player who copied the largest number of lines would receive €0.25 per line, whereas the last in the ranking would be paid €0.10 per line. Each player between the first and the last position would receive a per-line payment depending on his position such that the distance between two per-line remunerations is constant⁹. Under piece-rate payment, participants would receive €0.175 for each line copied correctly in the 45 minutes. The structure of the payments is such that the average value per line in the tournament is equal to the payment per line in the piece-rate scheme. Let us denote the tournament scheme with “T” and the piece-rate scheme with “PR”. The “job market” offered 146 positions (one for each experimental subject), of which one half allocated to T and the other half to PR. The participants were invited to bid for their preferred contract (either T or PR), knowing that winning either contract required falling in the highest quartile of the distribution of the bids, whereas the other participants would be randomly assigned T or PR, with a probability of 50 % and independently of their initially stated preference. The players expressed their bids as a percentage of their final payment, and were allowed to bid any amount between 0% and 100%. At the end of the experiment, the net payment for each participant was thus calculated as $(1 - \text{bid}) * \text{gross payment}$. As usual in auctions, only the winners had to pay their bids, whereas those who were randomly assigned a contract paid nothing. This mechanism allows evaluating the intensity of the each player’s preference for a given contract. This mechanism is not a standard auction, in the sense that it does not implement any of the well-known standard auction designs such as Dutch, all-pay, first or second price sealed-bid auctions, etc. However, the design used in this experiment was inspired by the extant literature on wage premiums for accepting unpreferred

contracts. This is typically observed in the case of (applicants) workers who accept fixed-term vs. open-ended contracts (see for example Pouliakis and Theodossiou, 2010). In such circumstances, the premium is expressed (and empirically estimated) in terms of percentage points of salary, since wages generally vary across several dimensions (what would render meaningless to provide absolute estimations). These wage premiums may also be interpreted as the “wage discount” that the workers are ready to accept to get an open-ended contract. The auction implemented in the experiment mirrors this situation; in other words, the subjects expressed the wage discount they are ready to accept to work under the preferred payment scheme (or, what is the same, the premium they require to work under the unpreferred contract). In this sense, the results of the implemented auction allow for assessing the preferences and their intensity for either auctioned contract¹⁰.

After the assignment of the contracts and after the participants were informed of their type, the task commenced. The assignment of contracts was a crucial element in this explorative study of gender-based preferences for and performance under competition. It allowed the analysis of whether players of a given gender prefer to engage in competition more than the other gender and whether competition enhances performance. It also reveals how players with a stated preference for competition, but who ended up with PR, performed in comparison to those who desired and received PR. Likewise, it allowed comparison of all players assigned T, whether or not they had a stated preference for competition. These last points also provide some indication as to the signalling value of the bids made during the auction: if ability and preference for competition correlate, this should show up in the results, with more capable individuals performing better - even in a non-competitive environment - than those who preferred to avoid competition.

Under the rules of the game, if people are rational and if they know their true relative ability, only $\frac{5}{12}$ ¹¹ of the participants should bid (a positive amount) for T. If this is not the case, then either some players are actually overconfident, or some have a misperception of their true relative position, or both.

The study presented here differs from Niederle and Vesterlund (2007) under some major aspects; firstly auctioning the contracts allows for evaluating how strong players' preferences are. Consequently it is possible to evaluate each player's degree of overconfidence (if any) more precisely than in the previous studies. Secondly the players can play only under the rules of one contract, hence their choice must be accurate, as they can not hedge as they can, to some extent, in Niederle and Vesterlund (2007). Notice that, when hedging is possible, players (and not only those who are overconfident) have more incentives to gamble than they have when hedging is not possible. Thirdly, if people are rational, their bids for the preferred contract should mirror their subjective expected position in the ability rank, under the veil of ignorance. We can also observe the behaviour of those who would have liked to compete, but who ended up playing under the piece-rate contract (and vice versa). This was, to some extent, possible also in Niederle and Vesterlund (2007), but with some crucial differences: in Niederle and Vesterlund (2007) the players first play under a given rule, and only then are given the possibility of choosing a contract; in this paper the choice is made *ex ante* and can not be undone. Actually, people who go on the job market for their first time are not familiar with their relative position in the ability ranking, hence the procedure used in this paper mirrors the real world better than Niederle and Vesterlund (2007) and allows for more realistic insights about the behaviour of those who enter the job market for the first time. The fact that some players did not obtain their preferred contract allows for testing whether the preference for a given payment scheme reveals some information about the potential performance of the subjects. This can be verified by comparing the actual performance of those who obtained the contract of their choice to the performance of players who were assigned a contract they did not choose.

I would also like to compare the experimental design used in this work with three other designs, to highlight the methodological strengths of my experiment and to explain why I chose the particular schemes used here. Freeman and Gelber (2010) pay subjects to solve mazes; under a scheme the payoff does not depend on the number of mazes solved, under another only the best

solver is awarded a large unique prize, under the last multiple prizes differentiated according to performance are given. This last scheme maximises the total effort (i.e. the number of solved mazes). The experiment proposed in the present paper uses a tournament that graduates the prizes in accordance with the subject's position in the final rank. As such, my experiment is designed to maximise the effort of all the participants. Vandergrift et al. (2007) design an experiment where they compare the subjects under three treatments: a) a graduated tournament, b) a winner-take-all tournament, and c) a piece-rate payment. Their results show that the effect of the incentive is maximum under (b), but that the entrants in the tournament select so that the best prefer (a) to (b). In other words, graduated tournaments are better selectors than winner-take-all tournaments, while these latter extract more effort than the former. Cason et al. (2010) compare a winner-take-all with a proportional-payment tournament and find that the second performs better than the first in terms of selecting the entrants; in particular the proportional-payment scheme limits the entries of poor subjects without altering the performance of the strong¹².

In order to check whether men and women differ systematically in their ability when recopying fictitious names and marks, I ran an additional session of the experiment (involving 40 new subjects), where the subjects were required to recopy 40 names and the relative marks in 45 minutes. Under this treatment, the subjects were paid the same amount of money, if they accomplished the task. This is very easy to perform also very slowly (as the average number of names recopied by the other subjects is more than 80). This allows comparing the speed at which males and women recopy names without any pressure and without any incentive for the fastest.

4. Core results

The analysis of the data is based on the Mann-Whitney test; accordingly with the hypotheses, first I present the results on overconfidence, and then those on gender effect and the comparison between preferences (although some intersection is possible).

4.1. Results of the auction

No player bid zero; hence none was indifferent between the two contracts. However, a large proportion of the subjects offered less than 5% of their final payment. Although any attempt to fix a threshold between weak and strong preferences would be arbitrary (and to my knowledge the extant literature does not help to solve the problem), it appears reasonable to consider “weak” the preference of those who bid less than 5% of their final payoff to gain a given contract.

Table I reports the results of the auction for the contracts. The average bid to win PR is higher than that for T, however the difference is never significant, neither for the pool of subjects, nor for the gender-homogeneous sub-samples. One might argue that, since women are more risk averse than men, then this affects the distribution of bids. However two facts do not support this hypothesis. The first is that 50% of males and 53.8% of females (the difference is not significant) bid a positive amount for the tournament. The second is that the bids for this contract do not differ between males and females significantly¹³. Now, assuming that females are more risk averse than males would reinforce the claims of the paper.

4.2. Results of the task

Before comparing the outcomes of the different treatments, I compare the overall abilities of men and women, in order to check whether any detected difference could be ascribable to differences not induced by the treatments, but inherent the natural heterogeneity of the sample. This serves to check whether the task was appropriate for the goals of the paper. In particular, I perform two tests. First, I compare the overall performance of all the men with that of all the women, independently of the treatment. The figures are 96.61 recopied rows for males and 97.85 for females. The difference between the figures is not statistically significant ($p\text{-value} = 0.737$) and suggests that there was no difference in the abilities between the genders. Second, I use the control treatment (where the subjects had to recopy 40 rows in 45 minutes) to compute the average time (in seconds) needed to recopy a row. This amounted to 26.15 seconds for men and to 24.47 seconds for women. A t-test reveals that these figures are not statistically different from

each other at any conventional level ($p\text{-value} = 0.532$). This suggests that men and women recopy fictitious names and marks at the same pace. Taken together these figures allow for assessing that men and women are equally able to perform the task assigned in this experiment. Therefore, any gender-related difference is confidently attributable to the treatments and to how the two genders react to these treatments; moreover the task chosen is appropriate.

I analyse now how the two treatments influenced the performance of the subjects, paying particular attention to their gender. The figures presented in Table II suggest that the performance of the women is not reactive to monetary incentives. The upper half of Table II shows the performance of the subjects, for a given assigned contract, independently of the preference revealed in the auction. While males perform significantly better under the tournament than the piece-rate scheme (differences significant at 10%), this is not the case for the women. The lower half of the table presents the results for the players' performance according to their revealed preferences. It is noteworthy that the self-sorting of males was a good signal of their final performance (either this is due to differences in ability or in effort or to both): those who bid for T performed much better than those who bid for PR (difference significant at 1% level) independently of working either under T or under PR. The same does not hold for women: here the difference between the performances is very tiny (about $\frac{1}{3}$ of men's) and not significant. Also, the women who preferred PR to T performed better than the men who did the same (although the difference is hardly significant: 10% level).

Table III shows the results of the two treatments, combining the information about the preferred and the assigned contract. The figures show mainly that, while the men who chose to work under the PR scheme took the right decision for maximising their payoff, this is not true for women, who worked hard under both. Women who preferred the PR contract worked always harder than males who expressed the same preference (third fourth of Table III, result significant at 5% level). This confirms once more that women are less reactive than men to monetary incentives. The fourth section of the table compares the losers of the auctions, i.e. the two sub-

samples of subjects who did not obtain the preferred contract, whereas the other parts of the table compare the subjects who bid for a contract and obtained it with those who bid for the same contract and did not obtain it. These figures further confirm the previous results: apparently women (try to) work hard, no matter which payment scheme they are assigned.

Another interesting result that emerges from Table III is that the men who bid for the tournament and worked under the piece-rate contract were more productive than the men who bid for the piece-rate scheme and worked under the tournament. Unfortunately, this result is weakened by the fact that part of the performance of the subjects who did not obtain the preferred contract may arise from sentiments such as envy for those who won the auction or frustration for losing it. Summarising the results reported so far, we observe that: on the one hand the women's average product is statistically the same, independently of the contract assigned and of whether this matches their preference or not. On the other hand, the men's product is higher under T than under PR and is higher for the men who bid for T than for the men who bid for PR.

5. Further results and discussion

The results reported in the previous section suggest that the exposure to competition (i.e. making the individuals work under a tournament scheme) is a useful tool to obtain high performances from male candidates, either because the best of them self-select in the tournament, or because males are very responsive to monetary incentives (or both). The last result mentioned may also suggest that men self-select between the two contracts according to their ability (thus revealing it to the observer), regardless effort. However, the results presented here cannot be taken as conclusive and further research is needed to confirm them. Women react differently than men to monetary incentives and contract schemes: indeed these do not induce self-selection between candidates nor render the women's performance higher under T than under PR. The women who were treated with the tournament always produced more than the colleagues who worked under the piece-ratescheme, whether they obtained the preferred contract or not. However, the

differences for the female sub-sample are never statistically significant and, therefore, we cannot conclude that monetary incentives foster the productivity of the women. In sum, monetary incentives are effective for males in making them to reveal their type (ability) and/or to foster their productivity, but fail the goal with females.

The players randomly assigned to the contracts are those who expressed a weak preference for either contract. One might argue that this selection can bias the results presented in Table III; Table IV tries to solve this possible objection. The figures reported in Table IV are analogous to those in Table II, but here I divide the sample according to the strength of the preference expressed by the participants. In particular, as mentioned before, strong preferences are represented by bids higher than 5% of the final retribution, while those between 0% and 5% (included) are considered weak¹⁴. The data are qualitatively the same in both sub-samples and confirm the previous findings: while, on average, the males with strong preferences for T performed better than those with weak preferences for the same contract (the figures are significantly different at 5% level), the ability of self-selecting¹⁵ is evident and significant in both the sub-groups. This reinforces the previous conclusion that males respond to monetary incentives either self-selecting between the two payment schemes or working harder to increase the payoff (or both). In either case the preference expressed through the bid is representative of their performance and provides (potential) employers with relevant information. This is not the case for women: not only the self-selection mechanism does not work, but the difference in performance¹⁶ shrinks as preferences get stronger (from 6.91 to 2.30 recopied lines), while for males the opposite holds (from 11.45 to 12.16 recopied lines).

Eventually Figures 1 and 2 show the correlation between the performance and the bid for winning the tournament: positive figures for this variable indicate the bids for T, while those for PR are represented as negative bids for T. It is noteworthy that, while for men the correlation is positive, for women it is negative (although the slope, in absolute terms, is steeper for males than for females). Indeed, one might expect low-ability players to prefer the piece-rate scheme and

high-ability players to prefer the tournament. Alternatively one might also expect that the subjects who bid for T put more effort and have a better performance than subjects who bid for PR. If such is the case, the correlation between the bid and the performance should be positive if the subject bids for T and otherwise negative (which is exactly what we observe for males). Of course the presence of some overconfidence can lead to weaker results than expected, but the observed reversal of the expected sign is a very strong result. This is especially true given that it holds for women, whom the literature tends to find less overconfident than men. The female subjects in this study display the opposite behaviour. This suggests that, whereas men are able to select the most profitable alternative for themselves, women are not (and they may even make choices that are not to their advantage).

Taken together, the results reinforce the previous conclusion that women are less sensitive to the incentive of competition than men: either their productivity remains unchanged in response to the incentive, or the incentive is not perceived as a stimulus to self-select according to their true abilities (corresponding to the likely scenario that they do not adjust their requests for compensation according to their true abilities), or both.

Another possible interpretation of the results is that women are more risk averse than men, and therefore they self-select between the two payment schemes more as a consequence of their risk preferences, rather than of their private information about their ability. The data used in this paper do not allow for ruling out this possibility, however if this is the reason why the self-selection between females is not elicited by monetary incentives (which are those mainly used in the labour market), it does not prevent the mechanism to fail in eliciting a signal from the female workers. And, in turn, this contributes to the existing gender pay gap. However a couple of results seem to suggest that differences in risk aversion (if any) play a minor role: on the one hand, the proportion of females who bid for the tournament is not significantly different from that of males; on the other hand, the average bid of women for T is larger than that of men¹⁷ (although the difference is very small and not statistically significant). Should there be some relevant difference

in the risk preferences of women on average, this should emerge from the data presented in Table I.

I would like to propose an additional reading of the results of the paper. Let us consider the well-known hypothesis of transitivity of preferences, such that if, given three alternatives (say A, B and C), a person prefers A to B and B to C, then transitivity entails that the same person prefers A to C. I would now apply this to the concept of productivity, as fostered by the two payment schemes presented in this experiment. We can observe that men's lines in contract T are significantly higher than men's lines in contract PR ($T_m > PR_m$). Men's lines in contract T are not significantly different from women's lines in contract T ($T_m = T_w$). Also men's lines in contract B are not significantly different from women's lines in contract PR ($PR_m = PR_w$). Then, (as $T_w = T_m > PR_m = PR_w$, this implies for transitivity that $T_w > PR_w$) women's lines in contract T should be significantly higher than women's lines in contract PR, but Table II reports that the last inequality does not hold¹⁸. Indeed, should the subjects be able to self-allocating between the two contracts in an efficient and effective way, then we should observe transitivity also in productivity. However, since this does not happen, we observe that of such a hypothesis is violated¹⁹.

6. Conclusions

One conclusion that can be drawn from the present data is that monetary incentives to productivity induce men either to signal their ability better than women do, or to work harder than men who are not incentivised (or both), while this is not the case for women. Moreover, women tend to accomplish the assigned task as well as they can, seemingly regardless of the incentive scheme. These results may help to explain the gender wage gap: women are not extremely sensitive to incentives, working hard always trying to maximise the payoff; of course this induces employers to incentivise (i.e. pay) them less, as the net marginal gain for the employer is much lower for a female than for a male worker²⁰. Moreover, women might also be less likely to or less interested in signalling their potential performance by asking for (or responding to) incentives so

that the employer can not use competitive contracts to select workers' types. Hence, they may tend to negotiate less with their employers than men. Finally, the results tend to suggest that women really do shy away from competition; the present data suggest that they do not perceive competition as a valuable incentive for working harder at their jobs. Of course this tendency is just one possible factor contributing to the wage gender gap, along with others such as discrimination, sexism, culture, preferences for child-caring and so forth. In addition, different levels of risk aversion between genders may play a relevant role. I have however to stress that monetary prizes are not the only incentive for workers; these may become more productive also in view of promotions, public praise and other "in-kind" prizes. And this might be true especially for women.

In this experiment as in the real life, ability is private information of a person. The results of the experiment presented in the paper do not allow for disentangling the effect of the monetary incentive on effort and that on self-selection between the two contracts. It is also noteworthy that also ex-post, ability and effort may be difficult to disentangle, as the observed performance is always a mix of both in proportions that are unknown to the observer. Nevertheless, this experiment points out that monetary incentives allow employers to extract more information and/or effort from male workers, but not from female workers.

The framework used here may mirror the selection procedure for young job candidates (including aptitude tests) or the working environment of fixed term workers, where both signalling and performance under a given scheme play a significant role²¹. However, the present data may help to explain the gender gap in wages also insofar as an employer earns a lower marginal return of incentives over women than over men. Gürtler and Kräkel (2010) suggest that the employer benefits from tournaments as these allow for extracting rents from the workers; however, this does not seem to hold when the workers are female. In other words, as the empirical evidence presented in this paper suggests, the employer extracts the maximum possible rent from women even in the absence of competitive incentives.

Assuming that competition in the workplace is stressful and should the women like competition less than men do (as found by Migheli, 2010 and 2014), the results of the paper suggest also another conclusion. Women perform equally under either contract; therefore, there is no reason why they should be “incentivised” by making them work in competitive environments. If both (or perhaps just either) the employers and (or) the women are (consciously) aware of this, then the rational choice for both (either) may be to choose a system of remuneration that does not entail competitive schemes. This would explain (at least partially) the wage gender gap, although it would not render the problem of gender discrimination on the workplace morally less important than it is.

References

- Bengtsson, Claes, Mats Persson and Peter Willenhag (2005). "Gender and Overconfidence" *Economics Letters*, 86(2): 199 – 203.
- Bertozzi, Elena (2008). "'You Play like a Girl!' Cross-Gender Competition and the Uneven Playing Field" *Convergence*, 14(4): 473 – 487.
- Biais, Bruno, Denis Hilton, Karine Mazurier and Sébastien Pouget (2005). "Judgemental Overconfidence: Self-Monitoring and Trading Performance in an Experimental Financial Market" *Review of Economic Studies*, 72(2): 287 – 312.
- Booth, Alison L. and Patrick J. Nolen (2012). "Choosing to Compete: how Different Are Girls and Boys?" *Journal of Economic Behavior & Organization*, 81(2): 542 – 555.
- Brooks, Peter and Horst Zank (2005). "Loss Aversion Behavior" *The Journal of Risk and Uncertainty*, 31(3): 301 – 325.
- Cason, Timothy N., William A. Masters and Roman M. Sheremeta (2010). "Entry into Winner-Take-All and Proportional-Prize Contests: an Experimental Study" *Journal of Public Economics*, 94(9-10): 604 – 611.
- Castagnetti, Carolina and Luisa Rosti (2009). "Effort Allocation in Tournaments: the Effect of Gender on Academic Performance in Italian Universities" *Economics of Education Review*, 28(3): 357 – 369.
- Cotton, Christopher, Frank McIntyre and Joseph Price (2010). "Gender Differences Disappear with Exposure to Competition" Working Paper no. 2010-11, University of Miami, Department of Economics.
- Croson, Rachel and Uri Gneezy (2009). "Gender Differences in Preferences" *Journal of Economic Literature*, 47(2): 448 – 474.
- Dargnies, Marie P. (2011). "Compétition en équipe: les homes sont-ils moins coopératifs que les femmes?" *Révue économique*, 62(1): 113 – 125.

- Dechenaux, Emmanuek, Dan Kovenock and Roman M. Sheremeta (2012). “A Survey of Experimental Research on Contests, All-Pay Auctions and Tournaments” Discussion Paper 2012-109, Social Science Research Centre Berlin.
- Dion, Paul A., Debbie Easterling and Raj Javalgi (1997). “Women in the Business-to-Business Salesforce” *Industrial Marketing Management*, 26(5): 447 – 457.
- Eckel, Catherine C. and Philip J. Grossman (2008). “Men, Women and Risk Aversion: Experimental Evidence” in (Charles Plott and Vernon Smith Eds.) *Handbook of Experimental Economics Results*, Amsterdam: Elsevier.
- Fehr-Duda, Helga, Manuele de Gennaro and Renate Schubert (2006). “Gender, Financial Risk, and Probability Weights” *Theory and Decision*, 60(2): 283 – 313.
- Fletschner, Diana, C. Leigh Anderson and Alison Cullen (2010). “Are Women as Likely to Take Risks and Compete? Behavioral Findings from Central Vietnam” *Journal of Development Studies* 46(8): 1459 – 1479.
- Freeman, Richard B. and Alexander M. Gelber (2010). “Prize Structure and Information in Tournaments: Experimental Evidence” *American Economic Journal: Applied Economics*, 2(1): 149 – 164.
- Gneezy, Uri, Muriel Niederle and Aldo Rustichini (2003). “Performance in Competitive Environments: Gender Differences” *The Quarterly Journal of Economics*, 118(3): 1049 – 1074.
- Gneezy, Uri and Aldo Rustichini (2004). “Gender and Competition at a Young Age” *American Economic Associations Papers and Proceedings*, 94(2): 377 – 381.
- Gneezy, Uri, Kenneth L. Leonard and John A. List (2009). “Gender Differences in Competition: Evidence from a Matrilineal and a Patriarchal Society” *Econometrica*, 77(5): 1637 – 1664.

- Gong, Binglin and Chun-Lei Yank (2012). “Gender Differences in Risk Attitudes: Field Experiments on the Matrilineal Mosuo and the Patriarchal Yi” *Journal of Economic Behavior & Organization*, 83(1): 59 – 65.
- Günther, Christina, Neslihan A. Ekinici, Christiane Schwieren and Martin Strobel (2010). “Women Can’t Jump? An Experiment on Competitive Attitudes and Stereotype Threat” *Journal of Economic Behavior & Organization*, 75(3): 395 – 401.
- Gürtler, Oliver and Matthias Kräkel (2010). “Optimal Tournament Contracts for Heterogeneous Workers” *Journal of Economic Behavior & Organization*, 75(2): 180 – 191.
- Ivanova-Stenzel, Radosveta and Dorothea Kübler (2011). “Gender Differences in Team Work and Team Competition” *Journal of Economic Psychology*, 32(5): 797 – 808.
- Kleinjans, Kristin J. (2009). “Do Gender Differences in Preferences for Competition Matter for Occupational Expectations?” *Journal of Economic Psychology*, 30(5): 701 – 710.
- Levitt, Steven D. and John A. List (2007). “What Do Laboratory Experiments Measuring Social Preferences Reveal about the Real World?” *Journal of Economic Perspectives*, 21(2): 153 – 174.
- Levitt, Steven D. and John A. List (2009). “Field Experiments in Economics: the Past, the Present and the Future” *European Economic Review*, 53(1): 1 – 18.
- Lindo, Jason M., Nicholas J. Sanders and Philip Oreopoulos (2010). “Ability, Gender, and Performance Standards: Evidence from Academic Probation” *American Economic Journal: Applied Economics*, 2(2): 95 – 117.
- Migheli, Matteo (2010). “Supporting the Free and Competitive Market in China and India: Differences and Evolution over Time” *Economic Systems*, 34(1): 73 – 90.
- Migheli, Matteo (2014). “Preferences for Government Interventions in the Economy: Does Gender Matter?” *International Review of Law and Economics*, 39: 39 – 48.

- Nekby, Lena, Peter Skogman Thoursie and Lars Vahtrik (2008). "Gender and Self-Selection into a Competitive Environment: Are Women more Overconfident than Men?" *Economics Letters*, 100(3): 405 – 407.
- Niederle, Muriel and Lise Vesterlund (2007). "Do Women Shy away from Competition? Do Men Compete too much?" *The Quarterly Journal of Economics*, 122(3): 1067 – 1101.
- Niederle Muriel and Lise Vesterlund (2008). "Gender Differences in Competition" *Negotiation Journal*, 24(4): 447 – 463.
- Ong, David and Zhuoqiong Chen (2013). "Tiger Women: an All-Pay Auction Experiment on the Gender Heuristic of the Desire to Win" (mimeo) available at SSRN: <http://ssrn.com/abstract=1976782>
- Plott, Charles R. (1982). "Industrial Organization Theory and Experimental Economics" *Journal of Economic Literature*, 20(4): 1485 – 1527.
- Price, Joseph (2008). "Gender Differences in the Response to Competition" *Industrial & Labor Relations Review*, 61(3): 320 – 333.
- Pouliakis, Konstantinos and Ioannis Theodossiou (2010). "Measuring the Utility Cost of Temporary Employment Contracts before Adaptation: a Conjoint Analysis Approach" *Economica*, 77(308): 688 – 709.
- Sabry, M. Abd-el-Fattah. (2010). "Longitudinal Effects of Pay Increase on Teachers' Job Satisfaction: a Motivational Perspective" *Journal of International Social Research*, 3(10): 11 – 21.
- Santos-Pinto, Luis (2012). "Self-Confidence: Wage Compression and the Gender Pay Gap" *Journal of Labor Economics*, 30(4): 873 – 914.
- Schwieren, Christiane and Doris Weichselbaumer (2010). "Does Competition Enhance Performance or Cheating? A Laboratory Experiment" *Journal of Economic Psychology*, 31(3): 241 – 253.

- Sjögren Lindquist, Gabriella and Jenny Säre-Söderbergh (2009). "Competitive; Successful and Conservative: Gender, Competition and Risk-Taking as Revealed in the TV Game-Show *Jeopardy*" Dissertation Series no. 59, The Swedish Institute of Social Research, University of Stockholm.
- Solnick, Sara J. (2001). "Gender Differences in the Ultimatum Game" *Economic Inquiry*, 39(2): 189 – 200.
- O'Toole, Daniel (2009). "Prize Structures, Gender, and Performance in Competitive Marathon Running" Haverford College, doctoral thesis.
- Vandergrift, Donald, Abdullah Yavas and Paul M. Brown (2007). "Incentive Effects and Overcrowding in Tournaments: an Experimental Analysis" *Experimental Economics*, 10(4): 345 – 368.
- Vandergrift, D. and Abdullah Yavas (2009). "Men, Women, and Competition: an Experimental Test of Behavior" *Journal of Economic Behavior & Organization*, 72(1): 554 – 570.

Table I. Bids over the two types of contracts. Standard deviations in brackets.				
	Average bid			Significance ¹
	Whole sample	Males	Females	
Bids for the preferred contract (either T or PR) independently of the assigned contract	9.68 (15.08)	8.74 (12.31)	10.55 (17.23)	-
Tournament	7.82 (8.33)	7.58 (6.90)	8.06 (9.65)	
Piece-rate	11.40 (19.23)	9.89 (16.03)	12.69 (21.72)	
Significance ²	-	-	-	
Sample size	146	69	77	

The figures represent the average percentage of the final remuneration that the subject is willing to pay to work under the preferred contract.

¹ The significance refers to the difference between the male and the female sub-samples. (figures in each row). Mann-Whitney test applied.

² This significance refers to the difference between the sub-samples working under the two different contracts (figures in each column).

Note: significance levels: *** (1%); ** (5%); * (10%); - non significant at any conventional level

Table II. Performance given the assigned or the preferred contract. S.d. in brackets				
	Average number of recopied lines			Significance ¹
	Whole sample	Males	Females	
<i>Lines copied under the assigned contract, independently of the preferred contract</i>				
assigned contract T	101.86 (24.30)	102.73 (27.05)	100.03 (21.71)	-
assigned contract PR	92.42 (19.18)	90.14 (20.67)	94.46 (17.78)	-
Significance ²	**	*	-	
<i>Lines copied under the preferred contract, independently of the assigned contract</i>				
preferred contract T	99.49 (21.37)	102.75 (20.92)	99.57 (22.08)	-
preferred contract PR	92.06 (22.58)	90.47 (27.08)	95.24 (17.91)	*
Significance ²	***	***	-	
Sample size	146	69	77	

The figures represent the average number of lines recopied. Mann-Whitney tests for differences between means.

¹ The significance refers to the difference between the male and the female sub-samples. (figures in each row).

² This significance refers to the difference between the sub-samples working under the two different contracts (figures in each column).

Note: significance levels: *** (1%); ** (5%); * (10%); - non significant at any conventional level

Table III. Performance given the preferred and the assigned contracts. Standard deviations in brackets

Preferred contract	Assigned contract	Average number of recopied lines			Significance ¹
		Whole sample	Males	Females	
T	T	99.44 (18.34)	99.04 (20.51)	99.81 (18.79)	-
T	PR	95.94 (12.08)	97.00 (10.55)	94.75 (14.26)	-
Significance ²		-	-	-	
Sample size		69	34	35	
PR	PR	88.87 (19.46)	83.62 (18.25)	93.07 (15.71)	**
T	T	99.44 (18.34)	99.04 (20.51)	99.81 (18.79)	-
Significance ²		***	***	-	
Sample size		77	35	42	
PR	PR	88.87 (19.46)	83.62 (18.25)	93.07 (15.71)	**
PR	T	91.42 (19.10)	82.87 (21.01)	97.64 (15.71)	**
Significance ²		-	-	-	
Sample size		73	32	41	
T	PR	95.94 (12.08)	97.00 (10.55)	94.75 (14.26)	-
PR	T	91.42 (19.10)	82.87 (21.01)	97.64 (15.71)	**
Significance ²		-	**	-	
Sample size		38	19	19	

The figures represent the average number of lines recopied. Mann-Whitney tests for differences between means.

¹ The significance refers to the difference between the male and the female sub-samples.

² This significance refers to the difference between the sub-samples working under the two different contracts (figures in each column).

Note: significance levels: *** (1%); ** (5%); * (10%); - non significant at any conventional level

Table IV. Performance given the assigned or the preferred contract. Subjects with weak or strong preferences. Standard deviations in brackets

	Average number of recopied lines			Significance ¹
	Whole sample	Males	Females	
Weak preferences (bid less than 5% of the final payoff)				
<i>Lines copied under the assigned contract, independently of the preferred contract</i>				
assigned contract T	96.18 (21.01)	91.93 (21.83)	99.32 (20.26)	-
assigned contract PR	91.33 (18.00)	89.95 (21.04)	92.41 (15.56)	-
Significance ²	-	-	-	
Sample size	91	39	52	
<i>Lines copied under the preferred contract, independently of the assigned contract</i>				
preferred contract T	99.55 (20.00)	98.26 (16.91)	100.61 (22.55)	-
preferred contract PR	90.86 (20.91)	86.81 (26.19)	93.70 (16.13)	-
Significance ²	**	***	-	
Sample size	91	39	52	
Strong preferences (bid more than 5% of the final payoff)				
<i>Lines copied under the assigned contract, independently of the preferred contract</i>				
assigned contract T	107.03 (25.41)	111.11 (27.42)	101.38 (22.12)	-
assigned contract PR	94.42 (21.43)	90.43 (20.85)	99.08 (22.05)	-
Significance ²	*	**	-	
Sample size	55	30	25	
<i>Lines copied under the preferred contract, independently of the assigned contract</i>				
preferred contract T	105.00 (23.14)	107.76 (24.18)	101.38 (22.12)	-
preferred contract PR	97.15 (20.34)	95.60 (18.37)	99.08 (22.05)	-
Significance ²	-	*	-	
Sample size	55	30	25	

The figures represent the average number of lines recopied. Mann-Whitney tests for differences between means.

¹ The significance refers to the difference between the male and the female sub-samples. (figures in each row).

² This significance refers to the difference between the sub-samples working under the two different contracts (figures in each column).

Note: significance levels: *** (1%); ** (5%); * (10%); - non significant at any conventional level

Figure 1. Preferences for the contracts and performance (females)

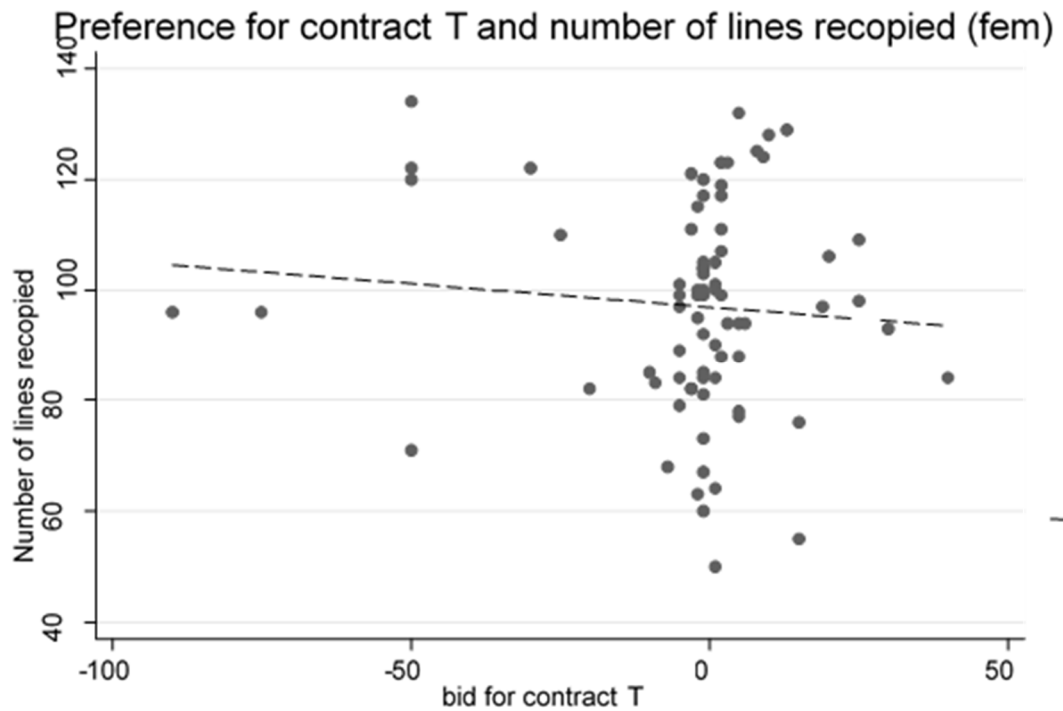
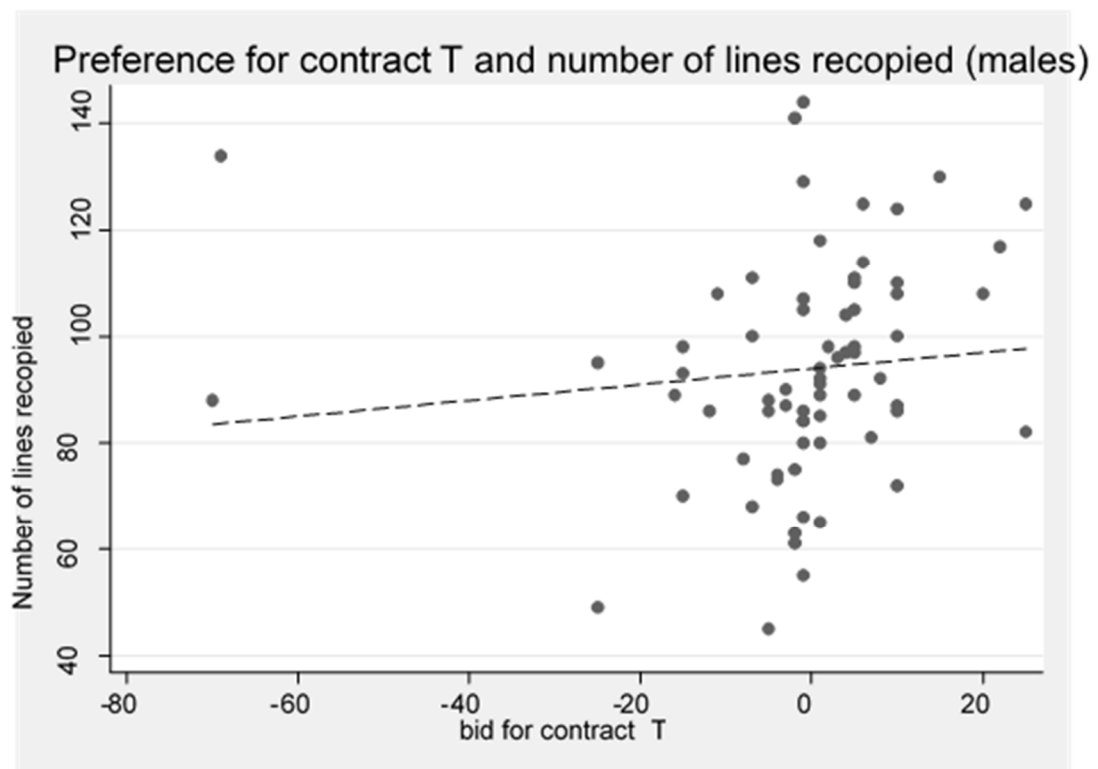


Figure 2. Preferences for the contracts and performance (males)



¹ It must be noted that in other contexts (such as in school) females usually perform better than males. However Lindo et al. (2010) find that academic probation at the end of the first year doubles

the probability of dropping out for males, but not for women. This evidence is in accordance with that in this paper. Assuming my results, indeed I can propose the following interpretation. Let's assume that there two types of students: of good (g) and of bad (b) quality. Now, when studying male students put an effort (E) which corresponds to their type; hence $E_{mg} > E_{mb}$. They do so, because they know that students of good type will find anyway jobs better remunerated than theirs. On the other hand, women do not respond to the monetary incentive in the job markets, but care for performing the best when assigned a task, independently of their type (even if they know their type). Hence, the difference $E_{fg} - E_{fb}$ should be lesser than the difference $E_{mg} - E_{mb}$, leading average higher marks for females than for males. Sabry (2010) finds that men's job satisfaction is positively affected by an increase in the salary, while women's is not. The author's results suggest that while men are more gratified than women by money, the latter are more gratified than the former by the attainment of a non monetary goal. Both the results of the economics of education literature and of my paper are in line with this.

² Nekby et al., (2008) show that (over)confidence pays off in terms of the results in competitive races; however, this result is not conclusive, as in some environments an excess of confidence can be detrimental for performance (Biais et al., 2005 and Sjögren Lindquist and Sæve-Söderbergh, 2009).

³ See also Günther et al. (2010), who find the same results but highlight that this happens only when the task is culturally viewed as a "male task". When this is culturally neutral (i.e., it is not perceived as "male" or "female"), competition increases the performance of both genders. Apparently women do not dislike competition *per se*, but dislike to compete against men.

⁴ However there could be some nurture effect that explains this result: Booth and Nolen (2012) find that women educated in all-female schools (where they are used to competing only against other females) are as competitive as men when examined in the framework of a field quasi-experiment,

but men are more competitive than women educated in mixed-gender schools, where they are used to also dealing with people of the opposite sex.

⁵ This means that, in this case, either competition is less important as a motivation, or the benefits from competing are offset by the composition of the team. In either case, this may explain why men tend to prefer individual competition to team-based competition (Dargnies, 2011).

⁶ Ivanova-Stenzel and Kübler (2011), p. 17.

⁷ See also Cotton et al. (2010).

⁸ Plott (1982), p. 1486.

⁹ Consider for example 5 players under the tournament scheme. The most performing would get €0.25 per line, the second in the ranking €0.2125, the third €0.175, the fourth €0.1375 and the last €0.10.

¹⁰ Since the paper focuses on the gender effect, I would like to spend some words on the possible consequences of using an auction mechanism. Women are generally found to be more risk averse than men are (Fehr-Duda et al., 2006 and Eckel and Grossman, 2008); therefore, we might expect that they are less competitive in auctions than men are. Nevertheless, the empirical evidence has found also the opposite result: women display more desire to win auctions than men do (Ong and Chen, 2013). Moreover the mediation of the pc may have attenuated the gender effect (Bertoizzi, 2008).

¹¹ The proof is the following: a total of 24 subjects participated in each session. In order to maximise the payoff, only the best half of them (i.e. 12 people) should bid for scheme T, as the median unit payment under scheme T is the same as the unit payment under scheme PR. However, the participant with median ability is indifferent between PR and T. For this reason, s/he should abstain from the auction, bidding 0. Therefore the proportion of people who should bid for T is $\frac{1}{2}$ -

$\frac{1}{12} = \frac{5}{12}$.

¹² In this section I aimed at presenting and discussing the closest and most relevant works to my design. However, there are several other works that use tournaments schemes in experimental studies. These are reviewed in Dechenaux et al. (2012).

¹³ Several works find that women have weaker preferences for competition than men are; therefore, the results presented in Table 1 may seem to be at odds with the extant literature. However, there are some possible explanations for this result. Bertozzi (2008) shows that when the interaction between the subjects is mediated by a pc or by social networks (as it is the case of my experiment) the differences of gender tend to disappear. Now, as women tend to compete less against men than against other women (Niederle and Vesterlund, 2008), the mediation of the pc may weaken the perception that women have about the gender composition of the group of competitors. This might render men and women more similar to each other in their attitude towards competition.

¹⁴ Indeed, as suggested by an anonymous referee, the least arbitrary method would be comparing the losers and the winners of the tournament (as I do in Table III). However the division between weak and strong preferences proposed in the paper serves as a robustness check of the results presented in Table III.

¹⁵ In this case, the “ability of self-selecting” refers to both the possibility of sorting according to ability and to that of responding to monetary incentives.

¹⁶ Measured as the difference between the lines recopied under T and those recopied under PRe for either strong or weak preferences.

¹⁷ Here one might argue that this is the signal that women are more risk averse: i.e. they bid more to increase the probability of winning the preferred contract. This can be true, but the differences with respect to males are small and not significant. Hence, if risk aversion does play a role, this is not very relevant.

¹⁸ I would like to thank an anonymous referee for suggesting this interpretation.

¹⁹ However, if we attenuate the strength of the inequalities and assume weak preferences (\geq) instead of strong preferences ($>$), the assumption of transitivity is not violated.

²⁰ Let p_m be the average productivity of men and p_f that of women; let a be the unit benefit for the employer, and let s be the amount of the incentive paid to the worker. Let p_i' , $i=m,f$ be the average productivity of category i after the introduction of the incentive. The profit for the employer be π before the introduction of the incentive and π' afterwards. We can write: $\pi_m = ap_m$ and $\pi_f = ap_f$ as well as $\pi'_m = ap'_m - s$ and $\pi'_f = ap'_f - s$. From these we can calculate the variations in the employer's profit in the case of each gender: $\Delta\pi_m = a(p'_m - p_m) - s$ and $\Delta\pi_f = a(p'_f - p_f) - s$, i.e. $\Delta\pi_m = a\Delta p_m - s$ and $\Delta\pi_f = a\Delta p_f - s$. Now, since the results of my experiment suggest that $\Delta p_m > 0$, whereas $\Delta p_f = 0$, it is clear that $\Delta\pi_m > \Delta\pi_f = 0$, which means that the employer has an incentive to stimulate men, but not women. A similar reasoning holds for the benefit gained by an employer who uses competition and ability-related wages as an incentive for job candidates to self-select according to ability.

²¹ However, as usual, the conclusions of an experiment are complex and difficult to generalize. It must also be said that 45 minutes of work in an experimental laboratory can not represent an entire career.